

STUDENT EVALUATION IN HIGHER EDUCATION: A COMPARISON BETWEEN COMPUTER ASSISTED ASSESSMENT AND TRADITIONAL EVALUATION

By

YARON GHILAY *

RUTH GHILAY **

* Lecturer in the Neri Bloomfield School of Design and Education, Haifa.

** Educational Counsellor in Primary Education.

ABSTRACT

The study examined advantages and disadvantages of computerised assessment compared to traditional evaluation. It was based on two samples of college students (n=54) being examined in computerised tests instead of paper-based exams. Students were asked to answer a questionnaire focused on test effectiveness, experience, flexibility and integrity. Concerning each characteristic, responders were asked to relate to both kinds of evaluation (computerised and traditional). Furthermore, students were asked to evaluate home and classroom computerised exams.

The research reveals that there is a significant advantage to computerised assessment in comparison to paper-based evaluation. The most powerful advantage of computer-assisted assessment found throughout the research, is a test's flexibility. The research findings point out that there is significant worthiness to adopt computerised assessment technologies in higher education, including home exams. Such a new method of evaluation is about to improve significantly the institutional educational administration.

Keywords: Computerised Assessment, Traditional Evaluation, Classroom Computerised Exams, Home Computerised Exams, Test Flexibility, Computer Assisted Assessment.

INTRODUCTION

The Department of Management at the Neri Bloomfield School of Design and Education, prepares students to teach management and accounting at high schools. The department's pedagogical aims are to provide students relevant tools, so they would be able to deal effectively with needs existing at high schools. The department deals with different levels, including theoretical and technological knowledge.

In the year 2009-10, a new Computer Assisted Assessment (CAA) system has been, firstly, used. The system, which is a part of the existing LMS (Learning Management System), has intended to replace traditional assessment. The first experiment of the new system was undertaken in the Department of Management, including the following courses

- Strategic management (fourth year).
- Entrepreneurship (fourth year).
- Scientific and technological literacy (third year).

- Statistical analysis via SPSS (third year).

In the year 2010-11, the new system was examined again including the same courses, except "scientific and technological literacy." This course has been replaced by another one - "management of technology" (third year).

In order to examine the effectiveness of the computerised tests, a research question was worded focused on the advantages and disadvantages of computer-assisted assessment in comparison to traditional evaluation. The intention was to gain general conclusions concerning the differences between computerised and paper-based exams, according to attitudes of students in a teacher-training college.

General Background

Assessment is a critical catalyst for student learning (Brown, Bull & Pendlebury, 1997) and there is considerable pressure on higher-education institutions to measure learning outcomes more formally (Farrer, 2002; Laurillard, 2002). This has been interpreted as a demand for more frequent

assessment. The potential for Information and Communications Technology (ICT) to automate aspects of learning and teaching is widely acknowledged, although promised productivity benefits have been slow to appear (Conole, 2004; Conole & Dyke, 2004). Computer Assisted Assessment (CAA), has a considerable potential both to ease the assessment load and provide innovative and powerful modes of assessment (Brown et al., 1997; Bull & McKenna, 2004), and as the use of ICT increases there may be 'inherent difficulties in teaching and learning online and assessing on paper' (Bull, 2001; Bennett, 2002a). CAA is a common term to the use of computers in the assessment of student learning. The term encompasses the use of computers to deliver, mark and analyse assignments or examinations. It also includes the collaboration and analysis of optically captured data gathered from machines such as Optical Mark Readers (OMR). An additional term is 'Computer Based Assessment' (CBA), which refers to an assessment in which the questions or tasks are delivered to a student via a computer terminal. Other terms used to describe CAA activities include computer based testing, computerised assessment, computer aided assessment and web based assessment. The term screen based assessment encompasses both web based and computer based assessment (Bull & McKenna, 2004).

The most common format for items delivered by CAA is objective test questions (such as multiple-choice or true/false) which require a student to choose or provide a response to a question whose correct answer is predetermined. However, there are other types of questions, which can be used with CAA.

CAA can also provide academic staff with rapid feedback about their students' performance. Assessments which are marked automatically can offer immediate and evaluative statistical analysis allowing academics to assess quickly whether their students have understood the material being taught, both at an individual and group level. If students have misconceptions about a particular theory/concept or gaps in their knowledge, these can be identified and addressed before the course or module's end.

Comparisons of Traditional Evaluation and CAA

The format of an assessment affects validity, reliability and student performance. Paper and online assessments may differ in several respects. Studies have compared paper-based assessments with computer-based assessments to explore this (Ward, Frederiksen & Carlson, 1980; Outtz, 1998; Fiddes, Korabinski, McGuire, Youngson & McMillan, 2002). In particular, the Pass-IT project has conducted a large-scale study of schools and colleges in Scotland, across a range of subject areas and levels (Ashton, Schofield & Woodgar, 2003; Ashton, Beavers, Schofield & Youngson, 2004). Findings vary according to the item type, subject area and level. Potential causes of mode effect include the attributes of the examinees, the nature of the items, item ordering, local item dependency and the test-taking experience of the student. Additionally there may be cognitive differences and different test-taking strategies adopted for each mode. Understanding these issues is important for developing strategies for item development as well as to produce guidelines for developing appropriate administrative procedures or statistically adjusting item parameters.

Limitations and Advantages of CAA

In contrast to marking essays, marking objective test scripts is a simple repetitive task, and researchers are exploring methods of automating assessment. Objective testing is now well established in the United States and elsewhere for standardized testing in schools, colleges, professional entrance examinations and for psychological testing (Bennett, 2002b; Hambrick, 2002).

The limitations of item types are an ongoing issue. A major concern related with the nature of objective tests is whether Multiple-Choice Questions (MCQs) are really suitable for assessing higher-order learning outcomes in higher-education students (Pritchett, 1999; Davies, 2002), and this is reflected in the opinions of both academics and quality assurance staff (Bull, 1999; Warburton & Conole, 2003). The most optimistic view is that item-based testing may be appropriate for examining the full range of learning outcomes in undergraduates and postgraduates, provided sufficient care is taken in their construction (Farthing & McPhee, 1999; Duke-Williams & King, 2001).

MCQs and multiple response questions are still the most frequently used question types (Boyle, Hutchison, O'Hare & Patterson, 2002; Warburton & Conole, 2003) but there is steady pressure through the use of 'more sophisticated' question types (Davies, 2001). Work is also being conducted during the development of computer-generated items (Mills, Potenza, Fremer & Ward, 2002). This includes the development of item templates precise enough to enable the computer to generate parallel items that do not need to be individually calibrated. Research suggests that some subject areas are easier to replicate than others—lower level mathematics, for example, in comparison with higher-level content domain areas.

Actually, CAA is not exactly a new approach. Over the last decade, it has been developing rapidly in terms of its integration into schools, universities and other institutions. Its educational and technical sophistication and its capacity to offer elements, such as simulations and multimedia-based questions, are not feasible with paper-based assessments (Bull & McKenna, 2004).

When there are increasing numbers of students and decreasing resources, objective tests may offer a valuable addition to existing ways of assessment, which are available for lecturers.

Possible advantages for using CAA might be the following

- To increase the frequency of assessment, thereby
Motivating students to learn.
Encouraging students to practice skills.
- To broaden the range of knowledge assessed.
- To increase feedback to students and lecturers.
- To extend the range of assessment methods.
- To increase objectivity and consistency.
- To decrease marking loads.
- To aid administrative efficiency.

(Bull & McKenna, 2004).

Method

The Research Questions

The research questions have been derived from the necessity to examine advantages and disadvantages of computerised assessment in comparison to traditional

evaluation (paper based exams) in higher education. Another aim was to examine if there are differences between home tests, and classroom computerised exams.

The following research questions were worded, relating to a teacher training college

- What are the advantages and disadvantages of CAA in comparison to traditional assessment methods, according to students' views?
- Are there advantages or disadvantages to computerised exams taken place at home in comparison to classroom tests, according to students' views?

Population and Samples

Population: The population addressed through the study included all students in the Neri Bloomfield School of Design and Education.

Samples: There were two samples included 54 students
Overall: 33 in the year 2010 and 21 in 2011. Students in the third and fourth year have been examined via Moodle computerised tests during the whole year. They were asked to answer a questionnaire at the end of the first semester of each academic year, concerning their perceptions towards computerised versus traditional exams.

The computerised exams related to the following courses (including open material)

- Strategic management (2010/2011).
- Entrepreneurship (2010/2011).
- Scientific and technological literacy (2010).
- Management of technology (2011)
- Statistical analysis via SPSS (2010/2011).

Each computerised exam included 25 multiple-choice questions with four or five answers each, except SPSS, which included different types of questions (multiple choice, calculated number and matching lists). Students were allowed to use any support material, and they had to finish the computerised exam during a definite time (110 minutes). When the time was over, the exam has been automatically submitted, having no chance to start over again.

The questionnaires were anonymous, and the rate of response was 90% (54 out of 60).

The traditional exams related to other courses existed in 2010/2011 (research methods, marketing, accounting, sociology, economics, psychology, management and organizational behaviour).

Tools

In order to examine the effectiveness of computerised learners' evaluation in comparison to traditional assessment, a questionnaire, including 48 closed questions was prepared: 24 items related to computerised assessment and 24 equivalent items to traditional one. The questionnaires were given to all the students who were examined in one computerised test at least. Most students were examined in two tests and some of them, took part in three or even four exams.

For each question, the respondents were requested to mention their views on the following Likert five-digit scale

- Strongly disagree
- Mostly disagree
- Moderately agree
- Mostly agree
- Strongly agree

The questionnaire was built based upon the literature review in order to identify the main variables relating to CAA. During the review, the following areas have been recognized as principal characteristics of CAA

- The frequency of evaluation.
- The level of coverage of knowledge areas being valued.
- Providing feedback.
- Diversity of methods and tools of evaluation.
- Objectivity and consistency.
- Workload linked to preparing, running and marking of exams.

In addition to the closed questions, the questionnaire included two open-ended questions as well. They were designated to accomplish the main data gathered by the quantitative principal part of the questionnaire, as follows

- Are there any additional strengths or weaknesses for computerised assessment beyond what has been mentioned earlier?
- Are there any additional strengths or weaknesses for paper-based assessment beyond what has been mentioned earlier?

Data Analysis

In order to examine the validity of the questionnaire, the reliability of the factors was calculated (Cronbach's alpha). Item analysis was undertaken as well in order to improve reliability. Based on the reliability found, the following 12 factors were built (2010 and 2011 together)

- *Test Effectiveness-CAA and Traditional Assessment:* Coverage of the taught material, accuracy, objectivity and consistency (two factors).
- *Test Experience-CAA and Traditional Assessment:* Convenience, pleasure/anxiety, ability of concentration, real time feedback (two factors).
- *Test Flexibility-CAA and Traditional Assessment:* Based on decreasing the load linked to preparation, transferring and marking, lots of opportunities, flexibility relating to dates for being examined (two factors).
- *Test Integrity-CAA and Traditional Assessment:* Accepting forbidden assistance, test questions leak, strictness on test discipline (two factors).
- *Satisfaction with Home and Classroom Computerised Tests:* Place preference, getting support from the lecturer, concentration, time convenience, flexibility, technical operation confidence, (two factors).
- *Test Integrity-Home and Classroom Computerised Tests:* Maintenance of test integrity, forbidden assistance, reflection of true knowledge (two factors).

For every single factor, there was found a high value of reliability (ranges from 0.649 to 0.891). Each factor has been determined by calculating the mean value of the items composing it.

Table 1 summarizes the eight factors (four for CAA and four for traditional assessment), the items composing them and reliability values.

Table 2 summarizes the other four factors (student

No. Factor	Questionnaire's questions
1 Test effectiveness: Computerised: Alpha=0.702 Traditional: Alpha=0.891	The test measures the level of my knowledge accurately. The test covers well the course material required. The test assesses basic learning objectives (knowledge and understanding). The test assesses high learning objectives (implementation, analysis, etc.). The test covers broad areas of the course. The test is objective and consistent.
2 Test experience: Computerised: Alpha=0.769 Traditional: Alpha=0.882	I enjoy the exam. I feel comfortable during the exam. The test score given at the end of the test is an advantage. I'm sure my answers would reach properly the lecturer. It is convenient for me to give answers on a computer screen. It is convenient to update answers I want to change prior to submission. I'm not worried about the exam. It is easy to concentrate while questions are displayed on a computer screen/paper. The test includes a variety of assessment methods. The time limit does not disturb me to concentrate on. I can appeal against examination results.
3 Test flexibility: Computerised: Alpha=0.649 Traditional: Alpha=0.889	I can get multiple opportunities to be tested. There are many opportunities to improve my grade. The lecturer can be flexible concerning the dates of exams.
4 Test integrity: Computerised: Alpha=0.670 Traditional: Alpha=0.853	It is difficult to get help from other examinees. There is no chance of a leak of exam questions. Examinees receive different test questionnaires. Test integrity is carefully maintained.

*Each question was written twice in the questionnaire – one for CAA and one for traditional assessment.

Table 1. Factors Relating to Computerised and Traditional Assessment, Including the Questionnaire's Questions *

satisfaction with computerised tests undertaken at home and in the classroom and test integrity relating to both places), the items composing them and reliability values.

The following statistical tests have been undertaken ($\alpha \leq 0.05$).

- *Independent Samples T-test:* It has been undertaken in order to check significant differences between each factor for the year 2010 in comparison to 2011.
- *Paired Samples T-test:* It was conducted for checking significant differences between computerised and traditional tests as well as additional pairs of factors.

No. Factor	Questionnaire's questions
1 Satisfaction with home/classroom computerised tests: Home - alpha=0.873 Classroom - alpha=0.878	I prefer to have a computerised test at home/classroom. A computerised test at home/classroom has an advantage over other examining alternatives. It is easy to get support from the lecturer during a computerised test at home/classroom. It is easy for me to concentrate on a home/classroom computerised test. A home/classroom computerised test allows me to be examined in convenient time. The flexibility of a home/classroom computerised test has a significant advantage. I feel confident concerning the technical operation of a home/classroom computerised test.
2 Computerised tests integrity-home/classroom: Home - Alpha=0.865 Classroom - alpha=0.840	Test integrity is carefully maintained during a computerised test at home/classroom. In a home/classroom computerised test, I do not receive assistance from others. In home/classroom computerised tests I get scores that reflect the true level of my knowledge.

*Each question was written twice in the questionnaire – one for home test and one for classroom test.

Table 2. Factors Relating to Student Satisfaction with Computerised Tests Undertaken at Home and in the Classroom and Test Integrity, including the Questionnaire's Questions*

Results

There was no significant difference between the years 2010 and 2011 concerning the mean scores of all questions and factors relating to both CAA and traditional assessment (ANOVA, $\alpha \leq 0.05$). It means that there was a replication of the results found in the first year (2010), also in the second year (2011). It strengthens the findings and gives them more validity. Mean factors' scores are presented for both years together in Table 3.

Factors	Mean	N	Std. Deviation	Significance of difference between computerised and traditional assessment
Test effectiveness - computerised	4.5462	52	.47104	$t_{(51)} = 3.276, P = 0.002$
Test effectiveness - traditional	4.1942	52	.73786	
Test experience - computerised	4.1173	53	.52112	
Test experience - traditional	3.7046	53	.81239	$t_{(52)} = 3.328, P = 0.002$
Test flexibility - computerised	4.3333	52	.57923	
Test flexibility - traditional	3.3782	52	1.16228	$t_{(51)} = 5.851, P = 0.000$
Test integrity - computerised	4.2010	51	.65004	
Test integrity - traditional	3.6650	51	.96256	$t_{(50)} = 3.999, P = 0.00$

Table 3. A comparison Between Computerised and Traditional Assessment

Table 3 shows that relating to these four factors, there is a significant advantage to CAA in comparison to traditional assessment.

Table 4 presents the gaps between all pairs introduced in Table 3. Comparison of these gaps shows that there is a significant difference between test flexibility (gap=1.17718) and all the three other gaps ($t_{(50)} = -3.777$, $p < 0.01$, $t_{(51)} = -3.444$, $p < 0.01$, $t_{(49)} = 2.666$, $p < 0.01$). On the other hand, there is no significant difference between the gaps relating to the other three factors. The meaning of these findings is that with regard to every single gap out of these four, there is a significant advantage to computerised tests in comparison to a traditional one. Further more, relating to tests' flexibility, the benefit of CAA is significantly greater in comparison to their advantage concerning the other three factors.

Table 5 presents a comparison between home and classroom computerised tests, regarding to satisfaction with the tests and the existing level of integrity. The findings show a significant advantage to satisfaction with home tests in comparison to classroom ones (both are computerised). However, relating to test integrity, there was no significant difference. Therefore, it can be confidently concluded that with regard to test integrity, home tests are at least not inferior in comparison to classroom exams.

Factors' mean gaps	N	Mean	Std. Deviation
Test flexibility	52	.9551	1.17718
Test integrity	51	.5359	.95703
Test experience	53	.4127	.90288
Test effectiveness	52	.3519	.77464

*Factors' mean gaps are sorted in descending order.

Table 4. Factors' Gaps: Computerised Test Mean Scores Minus Traditional Test Mean Scores *

Factors	Mean	N	Std. Deviation	Significance of difference between computerised and traditional assessment
Satisfaction with home computerised tests	4.0388	54	.83652	$t_{(53)} = 2.242$, $p = 0.029$
Satisfaction with class computerised tests	3.6238	54	.91126	
Computerised tests integrity -Home	4.0303	54	1.06205	$t_{(53)} = 0.123$, $p = 0.903$
Computerised tests integrity -Classroom	4.0123	54	0.64659	

Table 5. A Comparison Between Home and Classroom Computerised Tests

The open-ended questions strengthened the closed ones as shown in the following quotes

"The computerised test has no weaknesses-all the questions are clear, accurate and understood. I have no complains whatsoever."

"I enjoyed the computerised tests and in my opinion, it is definitely preferred in comparison to paper-based exams. A computerised test is much more convenient and interesting. In my view, computerised exams have only advantages."

The results summarized in Tables 3-5, the answers to the open-ended questions and statistical significant tests, have been the basis for wording answers to the research questions, as detailed in the next sections.

The research questions were as follows

I. What are the advantages and disadvantages of CAA in comparison to traditional assessment methods, according to students' views?

The results show that in students' view, computerised assessment has a significant advantage in comparison to traditional one, concerning the following factors being examined

- *Test Flexibility*: Test flexibility is expressed by the number of opportunities available for being examined, including chances for improving grades, as well as the lecturer's ability to adjust personal test times. Concerning this characteristic, a computerised exam has a decisive advantage (statistically significant) in comparison to all the other factors' advantage (the gap between computerised assessment and traditional one is 1.17718). Relating to the other factors, the computerised exam has also a significant advantage, although its strength is lower.
- *Test Effectiveness*: This factor describes how the exam measures relevant knowledge accurately, the material coverage, evaluation of learning objectives and the objectivity and consistency of the test. Relating to this factor, the computerised exam has a significant advantage in comparison to a paper-based, and the gap between computerised assessment and traditional one is 0.77464.

- *Test Experience*: It relates to the convenience and enjoyment of students from the test, the amount of anxiety, ability to concentrate, influence of the time limit and the possibilities to appeal. Concerning this factor, the computerised exam has a significant advantage in comparison to a traditional one, and the gap is 0.90288.
- *Test Integrity*: This factor describes how personal honesty is kept including getting forbidden help, questions' leak and exam discipline. Relating to this factor, the computerised exam has a significant advantage in comparison to a traditional one, and the gap is 0.95703.

ii. Are there advantages or disadvantages to computerised exams taken place at home in comparison to classroom tests, according to students' views?

The results show that in students' view, there was no significant difference between home and classroom computerised exams. Two factors have been examined

- *Students' Satisfaction*: It is expressed by their preferred place, level of support given by the lecturer, ability to be concentrated, time convenience, flexibility and the level of confidence concerning the technical operation of the computerised test. Concerning this characteristic, students' satisfaction with home computerised exams is better than classroom tests. The meaning of this finding is that students have no difficulties to operate home tests alone and feel confident to receive distance help while needed.
- *Tests' Integrity*: It is expressed by the ability to maintain test integrity, the extent to which unauthorized assistance is given and the extent to which tests reflect real knowledge. One of the greatest concerns of home tests is a hypothetic fear of keeping test integrity. The findings show that home tests' integrity is well maintained, at least equally to classroom exams.

Discussion

The literature review points out many advantages of CAA in comparison to traditional assessment. These benefits are mainly focused on organisational and managerial factors. When evaluating in a computerised form, it is much easier

to prepare, transfer and mark tests. Therefore, it is possible to cover a lot of material while reducing the burden on faculty and administrative staff. The worthiness of adopting a new computerised system of evaluation depends on its reliability and the ability to assimilate the necessary technological knowledge among lecturers.

Assuming that there is a significant advantage to computerised assessment for institutions of higher education, another critical question arises. The question is whether in "customers' view," namely students, computerised assessment is appropriate or at least does not cause difficulties in comparison to usual assessment. As such, it was necessary to examine the properties of the two methods of assessment from the students' perspective, in order to learn whether a computerised assessment has inferiority or on the contrary, it is superior.

Since the organisational and the administrative advantages are clear, it was enough to conclude that computerised assessment has at least no disadvantage for the examinees, in order to make it worthwhile to adopt the new technology. The study shows that not only there is no disadvantage with respect to computerised assessment criteria variety, but it found out that according to students' perspectives, information technology has significant advantages for them. The highlight is expressed in the best possible service to students due to the great flexibility of the computer system. If so, the worthiness of adopting computerised assessment technology increases significantly, and that might be a great contribution to the educational administration process.

Another important conclusion resulting from the research is that it is feasible to transfer computerised tests, which allow use of open material, at the student home instead of in the academic institution. This method has distinct organisational and managerial advantages but has also an advantage from the students' perspective. It allows great flexibility to students in terms of test date as well as not having to reach the institution of higher education. According to the research results, transfer to home computerised exams, does neither involve any disadvantages, nor problems relating to tests' integrity.

References

- [1]. Ashton H. S., Schofield D. K., & Woodgar S. C. (2003). Piloting summative web assessment in secondary education. In *Proceedings of the 7th International Computer-Assisted Assessment Conference* (eds Christie J.). Loughborough University.
- [2]. Ashton H. S., Beavers C. E., Schofield D. K., & Youngson M.A. (2004). Informative reports-experiences from the Pass-IT project. In *Proceedings of the 8th International Computer-Assisted Assessment Conference* (eds Ashby M. & Wilson R.). Loughborough University, Loughborough.
- [3]. Bennett R. E. (2002a). Inexorable and inevitable: the continuing story of technology and assessment. *Journal of Technology, Learning and Assessment*, 1, 100-108.
- [4]. Bennett R. E. (2002b). *Using Electronic Assessment To Measure Student Performance*. The State Education Standard, National Association of State Boards of Education.
- [5]. Boyle A., Hutchison D., O'Hare D., & Patterson A. (2002). Item selection and application in higher education. In *6th International CAA Conference* (eds Danson M.). Loughborough University, Loughborough.
- [6]. Brown G., Bull J., & Pendlebury M. (1997). *Assessing Student Learning in Higher Education*. Routledge, London.
- [7]. Bull J. (1999). Update on the National TLTP3 Project The Implementation and Evaluation of Computer-assisted Assessment. In *3rd International CAA Conference* (eds Danson M.). Loughborough University, Loughborough.
- [8]. Bull J. (2001). *TLTP85 Implementation and Evaluation of Computer-Assisted Assessment: Final Report*.
- [9]. Bull J. & McKenna C. (2004). *Blueprint for Computer-Assisted Assessment*. Routledge Falmer, NY.
- [10]. Conole, G. (2004). *Report on the effectiveness of tools for e-learning*, report for the JISC commissioned Research Study on the Effectiveness of Resources, Tools and Support Services used by Practitioners in Designing and Delivering E-Learning Activities.
- [11]. Conole, G. & Dyke, M. (2004). What are the affordances of Information and communication technologies?, *ALT-J*, 12(2), 111-123.
- [12]. Davies P. (2001). CAA must be more than multiple-choice tests for it to be academically credible? In *5th International CAA Conference* (eds Danson M. & Eabry C.). Loughborough University, Loughborough.
- [13]. Davies P. (2002). There's no confidence in multiple-choice testing. In *6th International CAA Conference* (eds Danson M.). Loughborough University, Loughborough.
- [14]. Duke-Williams E., & King T. (2001). Using computer-aided assessment to test higher level learning outcomes. In *5th International CAA Conference* (eds Danson M. & Eabry C.). Loughborough University, Loughborough.
- [15]. Farrer S. (2002). *End short contract outrage*. MPs insist, Times Higher Education Supplement.
- [16]. Farthing D., & Mc Phee D. (1999). Multiple choice for honours-level students? In *3rd International CAA Conference* (eds Danson M.). Loughborough University, Loughborough.
- [17]. Fiddes D.J., Korabinski A. A., McGuire G. R., Youngson M. A. & Mc Millan D. (2002). Are mathematics exam results affected by the mode of delivery? *ALT-J*, 10(6), 1-9.
- [18]. Hambrick K. (2002). *Critical issues in online, large-scale assessment: An exploratory study to identify and refine issues*. Capella University, Minneapolis.
- [19]. Laurillard D. (2002). *Rethinking university teaching a conversational framework for the effective use of learning technologies*. RoutledgeFalmer, London.
- [20]. Mills C., Potenza M., Fremer J., & Ward C. (2002). *Computer-based testing-building the foundation for future assessment*. Lawrence Erlbaum Associates, New York.
- [21]. Outtz J. L. (1998). Testing medium, validity and test performance. In *Beyond multiple choice evaluating alternative to traditional testing for selection* (eds Hakel M. D.). Lawrence Erlbaum Associates, New York.
- [22]. Pritchett N. (1999). Effective question design. In *Computer-Assisted Assessment in Higher Education* (eds Brown S., Race P. & Bull J.). Kogan Page, London).
- [23]. Warburton W., & Conole G. (2003). CAA in UK HEIs: the state of the art. In *7th International CAA Conference* (eds Christie J.). University of Loughborough, Loughborough.
- [24]. Ward W. C., Frederiksen N., & Carlson, S. B. (1980). Construct validity of free response and machine-scorable forms of a test. *Journal of Educational Measurement*, 7(1), 11-29.

ABOUT THE AUTHORS

Yaron Ghilay, Ph.D., is a Lecturer in the Neri Bloomfield School of Design and Education, Haifa, Israel and the Jerusalem College. He is also a tutor in professional specialization in educational technology at the Mofet Institute in Tel-Aviv. Previously, he has worked in secondary and higher education. His current research interests are associated with educational technology, school effectiveness, assessment and evaluation and teacher training.



Ruth Ghilay, Ph.D., is an Educational Counsellor in Primary Education. Previously, she has worked in educational roles in the military and in secondary education. Her current research interests are associated with educational technology, school effectiveness, assessment and evaluation and career transitions.

